Chapter 24

Situation Assessment and Decision Making in Skilled Fighter Pilots

◆

Wayne L. Waag
Herbert H. Bell
United States Air Force Armstrong Laboratory

This chapter presents preliminary findings from an attempt to identify situation assessment and decision making processes that account for differences in observed mission performance among skilled fighter pilots. The impetus for the study came directly from the U.S. Air Force Chief of Staff. In 1991, he posed a number of questions concerning situation awareness (SA) for the F-15 fighter world. First, what is SA? Can it be objectively measured? Is SA learned, or does it represent a basic ability? In response to the question, "What is it?," a working group at the Air Staff produced the following operator's definition of SA: "a pilot's continuous perception of self and aircraft in relation to the dynamic environment of flight, threats, and mission, and the ability to forecast, then execute tasks based on that perception" (Carroll, 1992). Although definitions of SA in the research literature focus primarily on processes underlying the assessment and resulting knowledge of the situation (Endsley, 1988), our working definition also included forecasting, decision making, and task execution. From a naturalistic, as well as an operational Air Force perspective, SA is more than simply knowledge and understanding of the environment.

The Armstrong Laboratory subsequently initiated a research investigation that had three goals: first, to develop and validate tools for reliably measuring SA (Waag & Houck, 1994); second, to identify basic cognitive and psychomotor abilities that are associated with pilots judged to have good

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **1997**	2. REPORT TYPE **Book Chapter**	3. DATES COVERED **01-06-1991 to 01-06-1994**

4. TITLE AND SUBTITLE **Situation assessment and decision making in skilled fighter pilots**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **62202F**
6. AUTHOR(S) **Wayne Waag; Herbert Bell**	5d. PROJECT NUMBER **1123**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Armstrong Laboratory,Operations Training Division,6030 South Kent Street,Mesa,AZ,85212-6061**	8. PERFORMING ORGANIZATION REPORT NUMBER **AFRL; AFRL/RHA**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061**	10. SPONSOR/MONITOR’S ACRONYM(S) **AFRL; AFRL/RHA**
	11. SPONSOR/MONITOR’S REPORT NUMBER(S) **AFRL-RH-AZ-BC-1997-0001**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Published in Naturalistic Decision Making, C. E. Zsambok & G. Klein (Eds.) (pp 247-254). Mahwah NJ: Lawrence Erlbaum Associates.

14. ABSTRACT
This chapter presents preliminary findings from an attempt to identify situation assessment and decision making processes that account for differences in observed mission performance among skilled fighter pilots. The impetus for the study came directly from the US Air Force Chief of Staff. In 1991, he posed a number of questions concerning situation awareness (SA) for the F-15 fighter world. From a naturalistic, as well as an operational Air Force perspective, SA is more than simply knowledge and understanding of the environment.

15. SUBJECT TERMS
Naturalistic decision making; Decision Making processes; Situation assessment processes; Fighter pilots; Performance measurement; Flight simulation; Training environments; Situation awareness;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**8**	

SA (Carretta, Perry & Ree, 1996); and third, to determine whether SA can be learned, and if so, to identify areas where cost-effective training tools might be developed and employed (Waag, 1994).

To develop measurement tools, it was first necessary to identify and describe critical behavioral indicators of the fighter pilot's ability to maintain good SA and successfully complete his mission. Earlier, Houck, Whittaker, and Kendall (1993) conducted a cognitive task analysis of a typical F-15 air combat mission. The resulting analysis identified the significant types of decisions required of the flight members, the information required for making these decisions, and the observable activities the flight members performed to acquire this information.

Using the taxonomy, a number of SA rating scales were developed to measure SA in operational units. Data were collected on 238 mission-ready F-15 pilots. From these data, a composite measure of SA was derived that was found to be highly related to previous flight experience (Waag & Houck, 1994). These measures also served as a means of selecting a sample of pilots to participate in a simulation phase of the effort, in which performance was observed under realistic combat conditions. During this phase, simulated air combat mission scenarios were developed for assessing SA, and a variety of performance measures were gathered in an attempt to determine those underlying situation assessment and decision making processes that distinguish pilots with good SA.

METHOD

Subjects

Forty F-15 pilots, who were flight lead qualified, served as subjects. An additional 23 F-15 pilots served as wingmen.

Equipment

The Armstrong Laboratory multiship simulation facility was used for data collection. This facility permits pilots to fly realistic air combat missions in a multibogey, high threat environment. The facility consists of a number of independent simulations operating as part of a secure distributed simulation network. For this study, the facility consisted of two high fidelity F-15C simulators, two F-16 simulators used as adversary stations, an automated threat engagement system, and an exercise replay system. This local simulation network located in Arizona was connected to an air weapons controller simulator at Brooks Air Force Base, TX by a dedicated telephone line. Additional details concerning the basic simulation architecture and components are available in Platt and Crane (1993).

Scenario Design

The approach taken toward the measurement of SA was through scenario manipulation and observation of subsequent performance as recommended by Tenney, Adams, Pew, Huggins, and Rogers (1992). Other approaches such as the use of explicit probes (Endsley, 1988) were considered and rejected due to their lack of face validity for the study participants. Because we were using mission-ready F-15 crews, it seemed essential that we provided a simulation experience as realistic as possible. A week-long SA "evaluation" exercise was constructed that consisted of 9 sorties, with 4 engagements per sortie. Over the week, engagements increased in complexity in terms of numbers of adversaries, enemy tactics, lethality of ground threats, type of controller support, etc.

A typical engagement scenario is presented in Fig. 24.1. This depicts an air combat mission in which the objective of the two F-15s is to defend the home airfield. In this case, the attackers consist of two bombers accompanied by two fighters. The engagement begins at 80 nautical miles (nm) separation in which the fighters are flying at 20,000 ft and the bombers at 10,000 ft. At 35 nm, the fighters begin a corkscrew type of maneuver in which they rapidly descend to 3,500 ft. At this time, they drop off of the F-15's radar screen. Upon completion of the maneuver, the fighters will trail the bombers and will be at a much lower altitude. Although F-15s can easily continue tracking the bombers, it requires the crew to "predict" the actions of the fighters so that they may be quickly reacquired on radar. At 15 nm, the bombers do a hard right turn and descend to 2500 ft. At this time, the bombers momentarily drop off the radar screen. Because the range is very close (10–12 nm), this situation requires the crew to accurately "predict" the actions of the bombers and correctly use their radar so that they may be quickly reacquired. The problem is further complicated in that the bombers and fighters now "merge" in roughly the same airspace. If the fighters are ignored, then they can launch their weapons against the F-15s. If the F-15s "lock" their radar on the fighters, which is usually the case at this point, then the bombers can continue toward the airfield "untargeted." Once the fighters are engaged, it is very difficult to reacquire the bombers since they are low and will be flying away from the F-15s. If the F-15s fail to kill the fighters, the problem will only be compounded.

This example not only shows the approach taken toward the design of the mission scenarios, but also serves to illustrate our contention that SA is more than knowledge of the current situation. In naturalistic environments, situation assessment and decision making are viewed as tightly coupled and are often difficult to separate. For fighter pilots to be successful, they must not only be able to "build the big picture," but they must also translate that assessment into a decision to employ weapons. Often, the inability to make these critical employment decisions may lead to mission failure, despite a correct assessment of the situation. In the sample scenario, the key to success

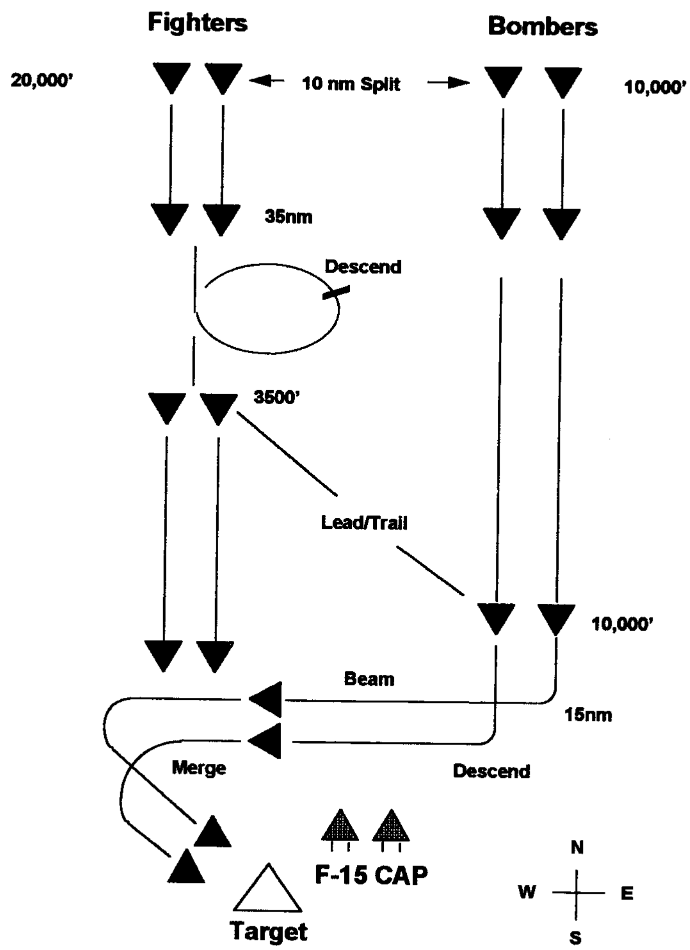

FIG. 24.1. Example engagement scenario.

is to target and destroy the bombers prior to 15 nm and then target the fighters. If the ranges become so close that all four threats must be dealt with simultaneously, then the mission is likely to fail. It is through the careful design of such mission scenarios that the failure to correctly assess the situation or to make correct employment decisions can be successfully inferred by observing pilot performance in the mission scenario.

Data Sources

In his discussion of process-tracing methodologies, Woods (1993) described the use of behavioral protocols. Essentially, behavioral protocols require that one gather as much data as possible so the investigator can piece together

the data in order to infer what underlying processes most likely occurred. Although a variety of objective data was gathered (Waag, 1994), the most important data sources, in our view, were the judgments and observations of two retired fighter pilots, who possessed an in-depth understanding of the air combat domain. The same two subject matter experts (SMEs) were used throughout the data collection effort. For each mission, one of the SMEs first attended the mission briefing session conducted by the crew. Both SMEs then observed each engagement and independently completed a checksheet to record pertinent events, notes, and outcomes. On completion of each mission, one of the SMEs accompanied the crew and observed the debriefing. The two SMEs then discussed each engagement and completed a consensus performance rating scale consisting of 24 behavioral indicators of SA related to F-15 mission performance. The SMEs also produced a written critical events analysis for each mission, which attempted to identify those events that, in their opinion, affected the outcome of the mission and were indicative of the crew's SA.

RESULTS

Given the voluminous amount of data gathered (1440 engagements), a major concern was how to select a subset of engagements that could be evaluated "in-depth" in an attempt to infer the processes that likely occurred. First, a composite score was computed for each pilot based upon the average performance rating assigned by the SMEs across all engagements. Based on this score, participants were rank ordered. The top 4 performing and bottom 4 performing pilots were then selected for exploratory analysis to identify whether or not there were consistent differences in certain classes of behaviors. The average number of flight hours in the F-15 for the top 4 was 1030 compared with the 719 for the bottom 4. All missions flown by each pilot were included.

The written critical event summaries were reviewed for the purpose of identifying comments indicative of both poor and good performance in hopes that they might differentiate the top performing and bottom performing pilots. Initially, five categories, loosely based on the taxonomy and the categories from the SA ratings scales, were defined—mission planning and tactics, system operation, communication, situation assessment, and weapons employment decisions. A sixth category, flight leadership, which represents the ability of the flight lead to effectively manage available resources, that is, the wingman and air weapons controller, was added due to the frequency of written comments by the observers.

Comments indicative of good performance tended to be fairly evaluative and nonspecific. The major discriminator between the top and bottom performing pilots was simply the frequency of positive comments (70 vs. 8). Frequency of comments falling in each category were then tallied. It is of

interest to note that the most frequently cited categories for the top performing pilots were flight leadership (23) and mission planning and tactics (15). Conversely, the least frequently cited categories were systems operations (0) and weapons employment decisions (1).

Comments indicative of poor performance were also placed in the six categories and frequencies were computed. These results are presented in Table 24.1. Several things should be noted. First, even the top group of pilots in the sample performed poorly at times. This is not surprising given the fact that the scenarios were designed to increase in complexity to a level with which the pilots were unfamiliar. Toward the end of the week, most of the engagements required the F-15 pilots to successfully defeat 6 enemy fighters plus 4 bombers. During operational training using the aircraft, such scenarios are extremely rare due to resource constraints. Second, and as expected, the bottom performing group produced about twice as many written comments indicative of poor performance.

Also of interest was the fact that neither of the groups produced any written comments by observers regarding system operations. In other words, at this level of competency, procedural operation of the onboard systems is mastered and not considered a problem. Of the six categories, weapons employment decisions clearly produced the highest number of "poor performance" comments, in fact, about twice as many compared with situation assessment. Usually, such comments focused on the failure to employ weapons, that is, take a shot, during the premerge phase of the engagement. Thus it would appear that weapons employment decisions may be more problematic than "building the big picture." It is also of interest to note that poor flight leadership comments were reported only for the bottom group. For the top rated performers, there were few comments regarding mission planning and tactics and none for flight leadership. For this group, poor weapons employment decision accounted for over half of the comments.

TABLE 24.1

Frequency of Poor Performance Comments Across Categories

Category	*Top Rated Performers*		*Bottom Rated Performers*	
	N	*%*	*N*	*%*
Planning and Tactics	8	22.8	17	27.4
Systems Operations	0	0	0	0
Communication	2	5.7	4	6.4
Situation Assessment	7	20	11	17.7
Employment Decisions	18	51.4	21	33.9
Flight Leadership	0	0	9	14.5
Totals	35	100	62	100

DISCUSSION

Although labeled differently, most cognitive models of information processing refer to at least three closely coupled and continuous components: perception and its meaning (situation assessment), decisions, and actions. For naturalistic settings, only actions are directly observable, the situational assessment and decision components must be inferred. Additionally, behavior in naturalistic environments is purposeful, and thus characterized by a structured goal hierarchy that gives direction to the behavior.

In the fighter environment, high level goals such as *destroy the enemy* and *avoid being destroyed* are fairly obvious. However, the mission plan, with tactics to be employed is less apparent and must also be inferred. It is often found that the mission plan is simply not working and that a change is necessary. The ability to successfully make such changes is a characteristic of the skilled fighter pilot. Again, it is the task of the observer to infer that such changes have occurred. And finally, the fighter pilot works in a team environment. The team in this case is very structured with flight lead in command of the other members including his wingman and air weapons controller. The ability to effectively lead must also be inferred by the observer, although in this instance the task is greatly simplified in the sense that verbal communication is the primary means whereby he exercises his leadership role.

The data gathered in the current study yielded some working hypotheses about the salience of each of these components, and how they change with the development of skill and proficiency. First, consider the basic assessment–decision–action cycle. The results suggest that task execution in terms of the actions involved in flying the aircraft and operating its systems is not a problem at the levels of competence in our sample. By the time the pilot is flight lead qualified, action has been mastered. Instead, most problems are manifested in the assessment and decision portion of the cycle. The results further suggest that problems involving employment decisions are twice as likely to be observed as problems involving situation assessment.

The results also point to the significance of decisions involving mission planning and tactics. Both groups produced comments indicative of tactical decisions that resulted in poor performance mission performance. These data, combined with data involving assessment and employment decisions, suggest that cognitive errors rather than procedural errors are the most likely cause of performance problems among competent fighter pilots.

The data also indicate that highly skilled performers are more likely to exercise effective flight leadership. As the flight leads' skill improves, they are better able to plan and adjust their tactics according to the demands of the scenario. They also improve in their ability to effectively manage all of the resources they have available. Thus, the most frequently cited comments regarding the top rated performers concerned their excellent flight leadership.

Taken as a whole, the results suggest that flight leadership, tactical mission planning, situation assessment, and decision making are critical for developing competent fighter pilots. The quality of such skills becomes manifested in the actions that lead to mission success. As Welford (1968) pointed out, these mental skills are critical components of skilled performance. From an operational perspective, the question becomes how to effectively and efficiently develop these skills. One approach is through multiship combat simulation such as those used in this study. In fact, the opinions of study participants were quite positive toward the use of simulation, with many citing its potential value for learning tactics and flight leadership skills. The value of such a medium lies in the fact that it provides real-world uncertainty which in turn forces the planning, assessment, decision, and action cycle within a naturalistic, team environment. Although mastery and practice of the individual component processes may provide some training value, it is our contention that such benefits will be limited. The development of competency in the fighter pilot community can be maximized to the extent that the training environment presents a realistic depiction of the richness and diversity of the naturalistic environment.

REFERENCES

Carretta, T. R., Perry, D. C., & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *International Journal of Aviation Psychology*, 6, 21–41.

Carroll, L. A. (1992). Desperately seeking SA. *TAC Attack* (TAC SP 127-1) 32: 5–6.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd annual meeting* (pp. 97–101). Santa Monica, CA: Human Factors Society.

Houck, M. R., Whittaker, L. A., & Kendall, R. R. (1993). An information processing classification of beyond-visual-range intercepts (Rep. AL/HR-TR-1993-0061). Brooks Air Force Base, TX: Armstrong Laboratory.

Platt, P., & Crane, P. (1993). Development, test, and evaluation of a multiship simulation system for air combat training. In *Proceedings of 15th Interservice/Industry Training Systems and Education Conference* (pp. 629–639). Orlando, FL: National Security Industrial Association.

Tenney, Y. J., Adams, J. J., Pew, R. W., Huggins, A. W. F., & Rogers, W. H. (1992). A principled approach to the measurement of situation awareness in commercial aviation (NASA Contractor Rep. 4451). Langley, VA: National Aeronautics and Space Administration.

Waag, W. L. (1994). Multiship simulation as a tool for measuring and training situation awareness. In *Proceedings of 16th Interservice/Industry Training Systems and Education Conference* (No. 5-3). Orlando, FL: National Security Industrial Association.

Waag, W. L., & Houck, M. R. (1994). Tools for assessing situational awareness in an operational fighter environment. *Aviation Space and Environmental Medicine*. 65(5, Suppl.)A13–19.

Welford, A. T. (1968). *Fundamentals of Skill*. London: Methuen & Co., Ltd.

Woods, D. D. (1993). Process-tracing methods for the study of cognition outside of the experimental psychology laboratory. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 228–251). Norwood, NJ: Ablex.